

**ANALISIS BUTIR SOAL PENILAIAN AKHIR TAHUN MATA PELAJARAN
BAHASA INDONESIA KELAS VIII SMPN 2 SUMBANG TAHUN PELAJARAN
2024/2025**

Apriliani¹, Memet Sudaryanto², Bivit Anggoro Prasetyo Nugroho³

^{1, 2, 3} PBI FIB Universitas Jenderal Soedirman

apriliani@mhs.unsoed.ac.id¹, memet.sudaryanto@unsoed.ac.id²,
bivit.nugroho@unsoed.ac.id³

ABSTRACT

The assessment aims to demonstrate the level of competency mastery of students and the effectiveness of the learning methods applied by educators. Based on interviews conducted with Indonesian language teachers of eighth grade students at SMP Negeri 2 Sumbang, it was found that the level of student achievement of the minimum passing grade (KKM) in the Final Year Assessment was still relatively low. The quality of the questions greatly determines the effectiveness of assessing student competency. This study aims to analyze the quality of the final assessment questions for the Indonesian language subject for eighth grade students at SMP Negeri 2 Sumbang for the 2024/2025 academic year. This study used a quantitative approach with a descriptive method. The research object consisted of 45 PAT questions with a sample of 144 students. Data were obtained through documentation techniques and analyzed using a classical test theory approach assisted by the Anates application. The analysis focused on the aspects of validity, reliability, difficulty level, discriminating power, and effectiveness of distractors. The results showed that the quality of PAT items still varied and did not fully meet the criteria for a quality assessment instrument. Some items were not valid, the test reliability coefficient was low, the difficulty level distribution was not balanced, the discriminating power of some items was still weak, and most distractors did not function optimally.

Keywords: *indonesian language, learning evaluation, question quality*

ABSTRAK

Penilaian bertujuan untuk menunjukkan secara konkret tingkat penguasaan kompetensi peserta didik serta efektivitas penerapan metode pembelajaran oleh pendidik. Berdasarkan hasil wawancara yang dilakukan dengan pendidik mata pelajaran Bahasa Indonesia kelas VIII di SMP Negeri 2 Sumbang, diperoleh temuan bahwa tingkat pencapaian peserta didik terhadap kriteria ketuntasan minimal (KKM) dalam pelaksanaan Penilaian Akhir Tahun masih tergolong rendah. Kualitas soal sangat menentukan efektivitas penilaian kompetensi peserta didik. Penelitian ini bertujuan untuk menganalisis kualitas butir soal Penilaian Akhir Tahun mata

pelajaran Bahasa Indonesia kelas VIII SMP Negeri 2 Sumbang Tahun Pelajaran 2024/2025. Penelitian menggunakan pendekatan kuantitatif dengan metode deskriptif. Objek penelitian berupa 45 butir soal PAT dengan sampel sebanyak 144 peserta didik. Data diperoleh melalui teknik dokumentasi dan dianalisis menggunakan pendekatan teori tes klasik berbantuan aplikasi Anates. Analisis difokuskan pada aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor. Hasil penelitian menunjukkan bahwa kualitas butir soal PAT masih bervariasi dan belum sepenuhnya memenuhi kriteria instrumen penilaian yang berkualitas. Sebagian butir soal belum valid, koefisien reliabilitas tes tergolong rendah, distribusi tingkat kesukaran belum seimbang, daya pembeda sebagian butir soal masih lemah, serta sebagian besar distraktor belum berfungsi secara optimal.

Kata Kunci: bahasa indonesia, evaluasi pembelajaran, kualitas soal

A. Pendahuluan

Penilaian merupakan penerapan berbagai alat penilaian dengan tujuan untuk memperoleh informasi mengenai hasil belajar peserta didik (Rangkuti & Albina, 2025). Melalui penilaian, pendidik dapat mengidentifikasi keunggulan dan kelemahan metode pembelajaran yang digunakan dalam pembelajaran. Dalam pelaksanaannya, penilaian membutuhkan sebuah instrumen penilaian yang baik dan sesuai dengan subjek yang akan diukur. Bentuk instrumen penilaian pembelajaran dapat berupa tes dan non tes tetapi dalam praktiknya guru lebih sering menggunakan tes sebagai instrumen penilaian karena lebih mudah dibandingkan nontes yang dianggap rumit (Sari & Yolanda, 2024).

Kurikulum merdeka membagi penilaian menjadi dua jenis utama, yaitu penilaian formatif dan penilaian sumatif. Abdullah & Hanifah (2025) mengungkapkan penilaian formatif bertujuan untuk memperbaiki dan meningkatkan proses pembelajaran, sedangkan penilaian sumatif berfokus pada evaluasi akhir untuk menggambarkan capaian hasil belajar peserta didik. Salah satu bentuk implementasi dari penilaian sumatif yaitu penilaian akhir tahun (PAT). Menurut Yuswardi dkk. (2023) penilaian akhir tahun adalah penilaian yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik pada akhir semester genap. Hasil dari PAT dapat digunakan oleh pendidik untuk bahan pertimbangan kebijakan pada raport peserta didik.

SMP Negeri 2 Sumbang merupakan sekolah menengah pertama yang berlokasi di Desa Banteran, Kecamatan Sumbang, Kabupaten Banyumas, Provinsi Jawa Tengah. Pelaksanaan penilaian akhir tahun (PAT) di SMP Negeri 2 Sumbang menggunakan bentuk tes yang terdiri atas kombinasi antara soal pilihan majemuk, soal benar-salah, dan soal uraian. Berdasarkan hasil wawancara yang dilakukan dengan pendidik mata pelajaran Bahasa Indonesia kelas VIII di SMP Negeri 2 Sumbang, diperoleh temuan bahwa tingkat pencapaian peserta didik terhadap kriteria ketuntasan minimal (KKM) dalam pelaksanaan PAT masih tergolong rendah. Terdapat kurang dari 1% peserta didik yang berhasil mencapai standar ketuntasan minimal yang ditetapkan oleh sekolah.

Pencapaian KKM yang rendah menandakan perlunya evaluasi ulang terhadap penilaian yang digunakan. Listyowati (2025) mengungkapkan kurangnya hasil penilaian berkaitan dengan kualitas soal yang digunakan. Hal ini sejalan dengan Firdausa dkk. (2025) yang menegaskan kualitas soal sangat menentukan efektivitas penilaian kompetensi peserta didik. Penentuan kualitas soal dapat

dilakukan melalui analisis butir soal. Menurut Auliya dkk. (2025), melalui analisis soal, dapat diperoleh informasi mengenai kualitas soal sehingga hasil tes lebih akurat dalam menentukan capaian kompetensi peserta didik. Analisis soal dapat dilakukan secara kuantitatif dengan fokus pada aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor.

Analisis terhadap soal penilaian yang akan diberikan kepada peserta didik diperlukan agar penilaian berkualitas sehingga hasilnya dapat mempresentasikan kemampuan peserta didik yang sebenarnya. Penelitian yang dilakukan Ashayidatina & Alber (2023) pada mata pelajaran Bahasa Indonesia kelas X di SMK N 1 Mempura bertujuan untuk menganalisis kualitas butir soal ujian semester ganjil dengan menggunakan analisis kuantitatif. Kelebihan dari penelitian tersebut yaitu data penelitian bersumber dari hasil ujian nyata peserta didik, sehingga temuan penelitian mencerminkan kondisi autentik pelaksanaan evaluasi pembelajaran di sekolah. Disamping kelebihan, penelitian ini juga memiliki keterbatasan yaitu tidak dilakukannya

analisis validitas dan reliabilitas instrumen tes. Penelitian hanya berfokus pada analisis tingkat kesukaran, daya beda, dan efektivitas pengecoh, tanpa menguji sejauh mana butir soal benar-benar mengukur kompetensi yang seharusnya diukur (validitas) serta konsistensi hasil pengukuran yang dihasilkan (reliabilitas).

Penelitian ini bertujuan untuk mengetahui kualitas butir soal Penilaian Akhir Tahun (PAT) mata pelajaran bahasa Indonesia kelas VIII di SMP Negeri 2 Sumbang Tahun Pelajaran 2024/2025 secara komprehensif. Kualitas butir soal dianalisis untuk memastikan bahwa instrumen penilaian yang digunakan benar-benar mampu mengukur kompetensi peserta didik sesuai dengan tujuan pembelajaran yang telah ditetapkan. Cakupan analisis pada penelitian ini lebih lengkap dibandingkan penelitian sebelumnya, yaitu dengan mengkaji validitas, reliabilitas, tingkat kesukaran, daya beda, serta efektivitas pengecoh secara terpadu dalam satu kajian. Berbeda dari penelitian sebelumnya yang umumnya hanya menelaah sebagian karakteristik empirik butir soal, penelitian ini menghadirkan

analisis menyeluruh terhadap seluruh aspek utama kualitas tes menggunakan pendekatan kuantitatif berbantuan aplikasi ANATES. Hasil penelitian ini tidak hanya memberikan gambaran objektif mengenai mutu soal PAT yang digunakan, tetapi juga menawarkan dasar empiris yang kuat bagi perbaikan, revisi, dan pengembangan bank soal Bahasa Indonesia di tingkat SMP secara berkelanjutan.

B. Metode Penelitian

Jenis penelitian yang digunakan dalam penelitian ini yaitu kuantitatif dengan menggunakan metode deskriptif. Hal yang akan di deskripsikan dan digambarkan pada penelitian ini yaitu mengenai kualitas dari butir soal Penilaian Akhir Tahun mata pelajaran bahasa Indonesia kelas VIII SMPN 2 Sumbang tahun pelajaran 2024/2025 yang ditinjau dari pendekatan klasik. Objek penelitian ini adalah soal PAT Bahasa Indonesia sebanyak 45 butir soal yang terdiri dari 35 butir soal pilihan majemuk, 5 butir soal benar salah, dan 5 butir soal uraian. Populasi penelitian mencakup seluruh peserta didik kelas VIII, dengan jumlah sampel sebanyak 144 peserta didik.

Teknik pengumpulan data yang digunakan yaitu dokumentasi. Teknik dokumentasi digunakan karena data yang diambil berupa dokumen-dokumen tertulis meliputi kis-kisi, kunci jawaban, soal, dan lembar jawaban peserta didik kelas VIII pada penilaian akhir tahun mata pelajaran Bahasa Indonesia SMP Negeri 2 Sumbang tahun pelajaran 2024/2025. Analisis data dilakukan secara kuantitatif dengan bantuan aplikasi anates untuk menentukan karakteristik empirik butir soal. Karakteristik empirik yang dianalisis meliputi validitas butir soal, reliabilitas tes, tingkat kesukaran, daya pembeda, serta efektivitas distraktor.

1. Validitas

Validitas adalah ketepatan suatu alat ukur dalam mengukur apa yang hendak diukur berdasarkan tujuan yang ingin dicapai. Pada software anates, validitas bisa dilihat pada korelasi skor butir dengan skor total yang dibagi menjadi dua kategori yaitu signifikan (soal valid) dan tanda (-) yang berarti tidak signifikan (soal tidak valid). Suatu soal dikatakan valid apabila r_{pb} lebih besar dari r_{tabel} . Nilai dari r_{tabel} diperoleh dari hasil taraf signifikansi 5% dengan n nya

sesuai dengan jumlah soal tes yang akan diteliti

2. Reliabilitas

Reliabilitas merupakan sejauh mana instrumen tersebut dapat dipercaya dan memberikan hasil yang konsisten. Soal memiliki reliabilitas tinggi jika koefisien realibilitas tes yang sedang diuji lebih besar dari 0,70 . Apabila lebih kecil dari 0,70 maka artinya tes yang diuji memiliki reliabilitas. Dalam aplikasi anates, analisis soal dilakukan sesuai dengan jenis soal. Sehingga hasil analisis reabilitas soal bukan dari keseluruhan soal melainkan dari masing-masing jenis soal.

3. Tingkat Kesukaran

Tingkat kesukaran soal berkaitan dengan seberapa banyak peserta didik dapat menjawab soal dengan benar. Hal ini sejalan dengan Manfaat dkk. (2021) yang menyatakan tingkat kesukaran dalam penilaian dilihat dari kemampuan peserta didik dalam menjawab soal, bukan dari asumsi guru yang menyusun soal, karena soal yang sulit /mudah bagi pendidik belum tentu sulit/mudah bagi peserta didik. Analisis tingkat kesulitan pada aplikasi Anates diklasifikasikan dalam lima kategori: sangat mudah, mudah, sedang, sukar,

dan sangat sukar (Khasanah dkk., 2023)

4. Daya Pembeda

Daya pembeda soal adalah kemampuan suatu butir soal dalam membedakan antara peserta didik berkemampuan tinggi dengan peserta didik berkemampuan rendah. Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi. Semakin tinggi indeks diskriminasi soal berarti semakin mampu soal yang bersangkutan membedakan siswa yang telah memahami materi dengan siswa yang belum memahami materi. Menurut Arikunto (2021), klasifikasi daya pembeda berdasarkan indeks daya pembeda yaitu baik sekali (0,71-1,00), baik (0,41-0,70), cukup (0,21-0,40), buruk (0,00-0,20), dan negatif (buruk sekali).

5. Efektivitas Distraktor

Menurut Arikunto (2021) sebuah distraktor dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut tes yang kurang memahami konsep. Menurut Toifur & Setyowati (2022), butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Namun sebaliknya, butir soal yang kurang baik

pengecohnya akan dipilih secara tidak merata oleh peserta didik. Adapun tanda keefektifan pengecoh pada anates berupa tanda (-) dan (+). Menurut Fiska dkk. (2021), pilihan jawaban pada butir soal berfungsi apabila bertanda (-) kurang baik, (+) baik dan (++) sangat baik, sedangkan bertanda (--) buruk, (---) sangat buruk dan terdapat blok yang menyatakan pilihan jawaban tidak berfungsi. Kategori efektivitas pengecoh pada setiap butir soal dibagi menjadi tiga yaitu kategori baik jika tiga pengecoh berfungsi, cukup jika dua pengecoh berfungsi, kurang baik jika hanya satu pengecoh yang berfungsi, dan sangat tidak baik jika semua pengecoh tidak berfungsi.

C. Hasil Penelitian dan Pembahasan

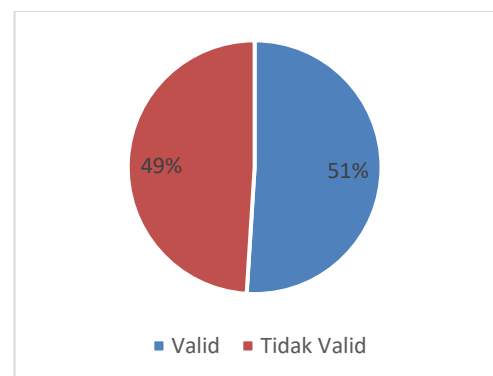
Analisis butir soal menjadi langkah penting untuk mengevaluasi penggunaan soal sehingga dapat menjadi dasar untuk perbaikan soal agar lebih baik (Ashayidatina & Alber, 2023). Analisis butir soal merupakan langkah penting dalam evaluasi kualitas instrumen penilaian karena memberikan informasi empiris mengenai ketepatan, konsistensi, dan keefektifan soal dalam mengukur kompetensi peserta didik.

Berdasarkan hasil analisis menggunakan bantuan aplikasi anates, kualitas butir soal Penilaian Akhir Tahun (PAT) mata pelajaran Bahasa Indonesia kelas VIII SMP Negeri 2 Sumbang Tahun Pelajaran 2024/2025 masih bervariasi dan belum sepenuhnya memenuhi kriteria sebagai instrumen penilaian yang berkualitas. Hasil analisis menunjukkan bahwa dari aspek validitas, sebagian butir soal masih tergolong tidak valid, yang mengindikasikan bahwa butir soal tersebut belum sepenuhnya mampu mengukur kompetensi yang sesuai dengan tujuan pembelajaran. Dari sisi reliabilitas, tes PAT menunjukkan koefisien reliabilitas yang rendah pada seluruh bentuk soal, sehingga konsistensi hasil pengukuran belum dapat diandalkan secara optimal. Analisis tingkat kesukaran memperlihatkan bahwa soal pilihan majemuk didominasi oleh kategori sedang, namun masih ditemukan proporsi soal yang terlalu mudah dan terlalu sukar. Pada aspek daya beda, sebagian besar butir soal berada pada kategori cukup hingga baik, meskipun masih terdapat sejumlah butir soal dengan daya beda buruk dan buruk sekali, khususnya pada bentuk pilihan

majemuk. Sementara itu, hasil analisis efektivitas distraktor menunjukkan bahwa sebagian besar pengecoh belum berfungsi dengan baik, namun masih ditemukan pengecoh yang efektif dan berfungsi.

1. Validitas

Berdasarkan hasil analisis validitas butir soal yang dikorelasi dengan skor total secara keseluruhan, dari 45 soal diperoleh 51% soal berkategori valid dan 49% soal berkategori tidak valid. Butir soal yang dikatakan valid yaitu nomor 2, 4, 7, 8, 9, 11, 13, 16, 19, 22, 23, 24, 25, 28, 35, 37, 38, 39, 40, 41, 42, 44, dan 45. Sedangkan untuk butir soal yang tidak valid yaitu nomor 1, 3, 5, 6, 10, 12, 14, 15, 17, 18, 20, 21, 26, 27, 29, 30, 31, 32, 33, 34, 36, dan 43. Temuan ini mengindikasikan bahwa hampir setengah butir soal belum mampu mengukur kompetensi yang seharusnya diukur sesuai dengan tujuan pembelajaran.



Gambar 1 Hasil Validitas Soal

Butir soal yang valid dapat mendukung tersedianya instrumen penilaian yang berkualitas, konsisten, dan efisien bagi pendidik. Sebaliknya, butir soal yang tergolong tidak valid perlu dilakukan revisi secara sistematis, baik dari segi kesesuaian dengan indikator, kejelasan rumusan, maupun ketepatan materi yang diukur, sebelum dapat digunakan kembali. Hal ini sejalan dengan penelitian yang dilakukan oleh Ramadhan dkk. (2024) yang menegaskan bahwa validitas berkaitan dengan ketepatan alat ukur terhadap konstruk yang diukur, sehingga soal yang tidak valid berpotensi menghasilkan informasi yang keliru tentang kemampuan peserta didik. Penggunaan butir soal yang tidak valid berpotensi menghasilkan data penilaian yang tidak akurat dan bias, sehingga instrumen penilaian tersebut tidak mampu mengukur kompetensi yang seharusnya diukur. Akibatnya, hasil penilaian tidak mencerminkan kemampuan peserta didik secara objektif dan dapat menimbulkan kesalahan dalam pengambilan keputusan pendidikan, seperti penentuan ketuntasan belajar maupun tindak lanjut pembelajaran.

2. Reliabilitas

Berdasarkan hasil analisis butir soal menggunakan aplikasi software Anates, reliabilitas soal Penilaian Akhir Tahun pada mata pelajaran bahasa Indonesia kelas VIII SMPN 2 Sumbang tahun pelajaran 2024/2025 menunjukkan koefisien reliabilitas yang rendah yaitu sebesar 0,49 untuk soal pilihan majemuk, 0,69 untuk soal benar salah, dan 0,52 untuk soal uraian. Suatu tes yang nilai reliabelnya rendah menandakan bahwa soal tes tersebut belum dapat dipercaya. Dengan kata lain, kompetensi yang diujikan pada peserta didik dalam penilaian ini belum atau tidak sesuai dengan kemampuan yang dimiliki oleh peserta didik. Berikut tabel hasil analisis reliabilitas pada soal penilaian akhir tahun bahasa Indonesia kelas VIII SMPN 2 Sumbang tahun pelajaran 2024/2025.

Tabel 1 Hasil Reabilitas Soal

Jenis soal	Reabilitas	Kategori
Pilihan majemuk	0,49	Rendah
Benar-salah	0,69	Rendah
Uraian	0,52	Rendah

Interpretasi koefisien reliabilitas secara umum menunjukkan bahwa nilai reliabilitas di bawah 0,70

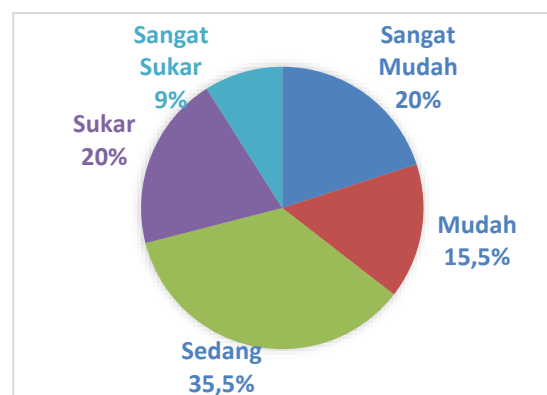
dikategorikan sebagai reliabilitas yang rendah atau kurang memadai sehingga skor yang dihasilkan instrumen belum dapat dipercaya sepenuhnya dalam mengukur kemampuan siswa secara konsisten dalam ranah kompetensi yang diuji. Anshari dkk. (2024) menyatakan bahwa reliabilitas tes mengacu pada konsistensi hasil yang diukur oleh suatu instrumen; tes yang reliabel secara statistik akan menghasilkan skor stabil ketika diuji berulang atau pada bentuk ujian yang sejenis, sementara tes dengan reliabilitas rendah cenderung menghasilkan variabilitas skor yang tinggi akibat faktor error, bukan perbedaan kemampuan nyata peserta didik.

Reliabilitas yang rendah tidak hanya mencerminkan konsistensi yang buruk, tetapi juga menurunkan kepercayaan terhadap kesimpulan evaluatif yang ditarik dari skor tes, seperti penilaian capaian kompetensi siswa atau pertimbangan keputusan pembelajaran lanjutan. Hal ini penting karena instrumen yang reliabel berkontribusi pada minimnya kesalahan pengukuran dan memberikan landasan data yang lebih stabil bagi pendidik untuk memetaka kemampuan peserta didik secara

akurat. hasil reliabilitas yang rendah pada ketiga bentuk soal tersebut menunjukkan bahwa instrumen evaluasi yang digunakan belum secara konsisten mencerminkan kemampuan peserta didik sesuai dengan kompetensi yang diujikan.

3. Tingkat Kesukaran

Berdasarkan hasil analisis butir soal menggunakan aplikasi Anates dan diklasifikasikan sesuai kategori tingkat kesukaran, dari total 45 butir soal ditemukan bahwa terdapat 20% butir soal dengan kategori sangat mudah, 15,4 % butir soal dengan kategori mudah, 35,5% butir soal dengan kategori sedang, 20% butir soal dengan kategori sukar, dan 9% butir soal dengan kategori sangat sukar. Berikut hasil tingkat kesukaran pada soal penilaian akhir tahun bahasa Indonesia kelas VIII SMPN 2 Sumbang tahun pelajaran 2024/2025.



Gambar 2 Hasil Tingkat Kesukaran Soal

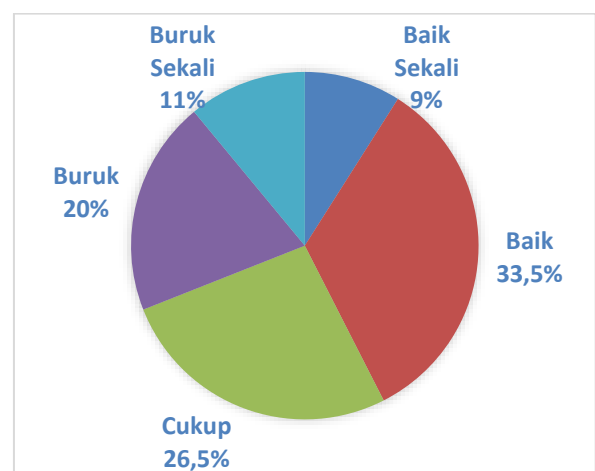
Suatu soal dapat dikatakan baik apabila adanya keseimbangan antara soal mudah dan soal sulit. Keseimbangan ini penting agar soal dapat mengukur kemampuan peserta didik dengan baik. Hal ini sejalan dengan Saputri dkk. (2023) yang menyatakan soal yang terlalu sulit dapat menyebabkan peserta didik bosan dan tidak berminat untuk belajar, sementara soal yang terlalu mudah membuat peserta didik tidak berusaha lebih maksimal untuk menyelesaikan soal. Soal yang baik harus dirancang agar mampu menantang kemampuan peserta didik tanpa tertekan.

Temuan distribusi tingkat kesukaran soal dalam penelitian ini menggambarkan perlunya perbaikan substansial dalam desain item, khususnya untuk menyeimbangkan proporsi soal-soal pada tingkat kesukaran sedang serta meminimalkan proporsi item yang ekstrem (sangat mudah atau sangat sukar). Hal ini penting supaya evaluasi pembelajaran benar-benar mampu mengukur kemampuan peserta didik secara akurat, adil, dan menyeluruh—memperhatikan variasi kemampuan siswa serta tujuan penilaian

pembelajaran Bahasa Indonesia dalam kurikulum.

4. Daya Pembeda

Hasil analisis daya pembeda butir soal yang dianalisis menggunakan software anates menunjukkan bahwa butir soal yang daya beda kategori sangat baik berjumlah 4 soal pada soal nomor 10, 15, 16, 18. Butir soal yang daya bedanya termasuk dalam kategori baik berjumlah 6 soal pada soal nomor 2, 4, 7, 8, 14, 19. Butir soal yang daya bedanya termasuk dalam kategori kurang baik berjumlah 5 soal pada soal nomor 3, 6, 11, 12, 20 dan butir soal yang daya bedanya termasuk dalam kategori tidak baik berjumlah 5 soal pada nomor 1, 5, 9, 13, 17. Berikut hasil daya beda pada soal penilaian akhir tahun bahasa Indonesia kelas VIII SMPN 2 Sumbang tahun pelajaran 2024/2025.



Gambar 2 Hasil Daya Pembeda Soal

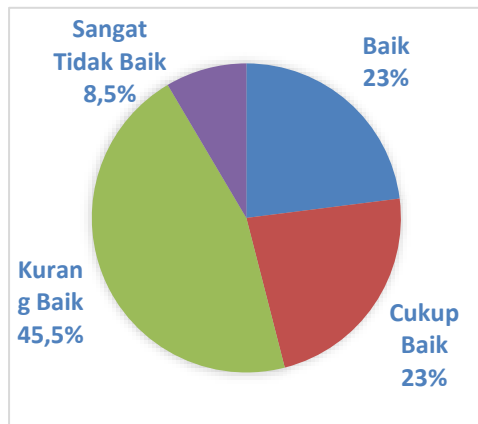
Daya pembeda merupakan indikator penting dalam analisis butir tes karena menunjukkan kemampuan suatu item untuk membedakan antara siswa dengan kemampuan tinggi dan rendah, sehingga item dengan daya pembeda tinggi dapat secara efektif memisahkan kelompok siswa berdasarkan ketercapaian kompetensi. Distribusi kategori daya pembeda dalam penelitian ini mengindikasikan bahwa sebagian item soal belum optimal dalam membedakan tingkat kemampuan siswa. Butir soal yang masuk dalam kategori kurang baik maupun tidak baik sering kali kurang sensitif dalam menangkap perbedaan kemampuan peserta didik sehingga dapat berdampak pada akurasi skor ujian dalam menggambarkan kompetensi nyata siswa. Butir-butir soal yang berada pada kategori kurang dan tidak baik sebaiknya direvisi untuk meningkatkan sensitivitasnya terhadap variasi kemampuan siswa, mengukur kompetensi bahasa Indonesia sebagaimana diharapkan dalam tujuan pembelajaran kurikuler.

5. Efektivitas Distraktor

Berdasarkan hasil analisis efektivitas distraktor butir soal Penilaian Akhir Tahun (PAT) mata

pelajaran Bahasa Indonesia kelas VIII SMPN 2 Sumbang Tahun Pelajaran 2024/2025 dilakukan menggunakan aplikasi Anates, diperoleh gambaran bahwa kualitas distraktor soal masih bervariasi. Dari keseluruhan 35 butir soal yang dianalisis, sebanyak 8 butir soal (23%) berada pada kategori baik, yang menunjukkan bahwa pengecoh pada soal-soal tersebut mampu berfungsi secara optimal karena dipilih oleh peserta didik dan dapat membedakan kemampuan siswa. Selanjutnya, 8 butir soal (23%) termasuk kategori cukup, artinya distraktor masih berfungsi meskipun efektivitasnya belum maksimal. Sementara itu, jumlah terbesar terdapat pada kategori kurang baik, yaitu 16 butir soal (45,5%), yang mengindikasikan bahwa sebagian besar distraktor kurang menarik bagi peserta didik atau tidak berfungsi sebagaimana mestinya. Adapun 3 butir soal (8,5%) masuk dalam kategori sangat tidak baik, yang menunjukkan bahwa distraktor pada soal-soal tersebut tidak dipilih sama sekali atau sangat lemah dalam mengecoh peserta didik. Berikut hasil analisis efektivitas Distraktor pada soal penilaian akhir tahun bahasa

Indonesia kelas VIII SMPN 2
Sumbang tahun pelajaran 2024/2025.



Gambar 3 Hasil Uji Efektivitas Distraktor Soal

Variasi efektivitas distraktor juga dapat mencerminkan kurangnya keselarasan antara indikator soal, materi pembelajaran, dan alternatif jawaban yang disusun. Distraktor yang disusun tanpa mengacu secara jelas pada indikator pencapaian kompetensi akan kehilangan daya tariknya bagi peserta didik, karena pilihan jawaban menjadi terlalu mudah dieliminasi. Akibatnya, peserta didik dapat menjawab soal dengan benar melalui proses eliminasi sederhana, bukan karena penguasaan konsep yang sesungguhnya. Keberadaan distraktor yang sangat lemah atau tidak dipilih sama sekali menunjukkan adanya kelemahan mendasar dalam konstruksi opsi jawaban, khususnya pada aspek homogenitas dan kejelasan struktur bahasa.

D. Kesimpulan

Penelitian ini menunjukkan bahwa kualitas butir soal Penilaian Akhir Tahun (PAT) mata pelajaran Bahasa Indonesia kelas VIII SMP Negeri 2 Sumbang Tahun Pelajaran 2024/2025 masih belum optimal sebagai instrumen evaluasi pembelajaran. Temuan utama memperlihatkan bahwa validitas butir soal belum merata, reliabilitas tes tergolong rendah, tingkat kesukaran belum terdistribusi secara seimbang, daya pembeda sebagian butir soal masih lemah, serta distraktor pada banyak soal belum berfungsi secara efektif. Kondisi tersebut menandakan bahwa instrumen penilaian belum sepenuhnya mampu mengukur kompetensi peserta didik secara akurat dan konsisten. Hasil penelitian ini menegaskan perlunya perbaikan berkelanjutan dalam penyusunan dan pengembangan soal melalui analisis butir secara empiris, peningkatan ketelitian dalam merumuskan indikator dan pilihan jawaban, serta penguatan kolaborasi antarpendidik dalam pengembangan bank soal. Dengan perbaikan tersebut, instrumen penilaian diharapkan dapat memberikan gambaran capaian belajar peserta didik secara lebih

objektif dan mendukung peningkatan kualitas pembelajaran Bahasa Indonesia di tingkat pendidikan menengah.

DAFTAR PUSTAKA

- Abdullah, M. S., & Hanifah. (2025). Model Evaluasi Formatif dan Sumatif: Strategi untuk Meningkatkan Proses dan Hasil Pembelajaran di Pendidikan Dasar pada Kurikulum Merdeka. *Jurnal Ilmiah Multidisiplin*, 2(4), 30–35.
<https://doi.org/10.62017/merdeka>
- Anshari, M. I., Nasution, R., Irsyad, M., Alifa, A. Z., & Zuhriyah, I. A. (2024). Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI. *EDUKATIF: JURNAL ILMU PENDIDIKAN*, 6(1), 964–975.
<https://doi.org/10.31004/edukatif.v6i1.5931>
- Arikunto, S. (2021). Dasar-dasar evaluasi pendidikan edisi 3. Jakarta: Bumi aksara.
- Ashayidatina, L., & Alber. (2023). Analisis Butir Soal Ujian Semester Ganjil Bahasa Indonesia Kelas X SMK Negeri 1 Mempura. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, 8, 131–137.
<https://dx.doi.org/10.26737/jp-psi.v8i2.5039>
- Auliya, A. S., Hikmah, N., Tyaningsih, R. Y., & Turmuzi, M. (2025). Analisis Kualitas Butir Soal Asesmen Tengah Semester

- Genap Mata Pelajaran Matematika Kelas VII SMP Negeri 2 Kuripan Tahun Pelajaran 2024/2025. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, 10(2), 210–224.
<https://doi.org/10.23969/jp.v10i02.25625>
- Firdausa, W. F., Delawati, D., & Yulianti. (2025). Analisis Butir Soal Pilihan Ganda Penilaian Akhir Semester pada Pembelajaran Tema 8 Praja Muda Karana Kelas III. *Jurnal Sadewa: Publikasi Ilmu Pendidikan, Pembelajaran Dan Ilmu Sosial*, 3(2), 129–151.
<https://doi.org/10.61132/sadewa.v3i2.1727>
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian IPA Menggunakan Software Anates pada Pendekatan Teori Klasik. *Jurnal Natural Science Educational Research*, 4(1), 65–76.
<https://doi.org/10.21107/nser.v4i1.8133>
- Khasanah, I., Fuady, A., & Sunismi. (2023). Analisis Soal Ulangan Harian Matematika Bentuk Pilihan Ganda. *MATHEMA JOURNAL*, 5(2), 110–125.
<https://doi.org/10.33365/jm.v5i2.2818>
- Listyowati, I. (2025). Analisis Butir Soal Multiple Choice Mata Pelajaran Bahasa Indonesia Siswa Kelas 4 di SDN Bareng 5 Jombang. *Jurnal Review Pendidikan Dan Pengajaran*, 8(2), 3968–3976.

- <https://doi.org/10.31004/jrpp.v8i2.44453>
- Manfaat, B., Nurazizah, A., & Misri, M. A. (2021). Analysis of mathematics test items quality for high school. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 25(1), 108–117.
<https://doi.org/10.21831/pep.v25i1.39174>
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas and Reliabilitas. *Journal on Education*, 06(02), 10967–10975.
<https://doi.org/10.31004/joe.v6i2.4885>
- Rangkuti, M. H., & Albina, M. (2025). Penilaian Dan Evaluasi Pembelajaran Dalam Meningkatkan Efektivitas Pembelajaran Pada Lembaga Pendidikan Islam. *Qosim : Jurnal Pendidikan, Sosial & Humaniora*, 3(1), 358–366.
<https://doi.org/10.61104/jq.v3i1.829>
- Saputri, H. A., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis Instrumen Assesmen : Validitas, Reliabilitas, Tingkat Kesukaran dan Daya Beda Butir Soal. *Didaktik : Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 9(5), 2986–2995.
<https://doi.org/10.36989/didaktik.v9i5.2268>
- Sari, A. P. I., & Yolanda. (2024). Validitas dan Reliabilitas Soal Ujian Sumatif Tengah Semester (STS) Kelas X Fase-E SMA Negeri 8 Palembang Tahun Pelajaran 2023/2024. *ALFABETA: Jurnal Bahasa, Sastra, Dan Pembelajarannya*, 7(2), 337–347.
<https://doi.org/10.33503/alfabeta.v7i2.134>
- Toifur, I., & Setyowati, E. (2022). Analisis Butir Soal Simulasi Ujian Sekolah Mata Pelajaran Matematika SMA Negeri 1 Salaman. *MATH LOCUS: Jurnal Riset Dan Inovasi Pendidikan Matematika*, 3(1), 28–39.
<https://doi.org/10.31002/mathlocus.v3i1.2341>
- Yuswardi, Nawawi, E., & Hestiriniah, D. C. (2023). Analisis Pelaksanaan Penilaian Akhir Tahun Di Sekolah Dasar Negeri kecamatan Indralaya Utara Kabupaten Ogan Ilir. *@-Publik Jurnal Administrasi Publik*, 3(3).
<https://doi.org/10.37858/publik.v3i3.402>